

全国第六届研究生数学建模竞赛



题 目 基于一阶不变矩方法的弹痕自动比对研究

摘 要：

本文主要利用一阶不变矩提取两颗子弹擦痕特征，计算擦痕相似度和子弹相似度，从而判断两颗子弹是否来自于同一支枪。

首先，对擦痕测量数据进行分析，根据仪器在 z 轴方向上测量精度对 z 方向上的测量数据进行有效数字处理。然后在数据空间域使用邻域平均的方法来平滑数据、去除数据误差和随机噪声，并给出了数据误差及噪声消除前后的图像对比。

考虑到不同擦痕测量数据没有统一的参考系，于是从测量误差的来源出发，提出了利用弹头自身测量数据调整自身位置和姿态的自校准方法，从而使弹头轴线尽可能与基准平面平行，使擦痕走向尽可能与 y 轴平行。这样就尽可能保证两个弹头在相同的位置和姿态下进行比对。

擦痕特征的提取是本文重点研究的内容。将擦痕理解为图像，则擦痕的比对问题可转化为图像的特征提取和比对问题。本文利用一阶不变矩来描述数据的局部特征。该方法的优点是提取擦痕特征不受擦痕数据空间坐标变换的影响。通过对需要比对的两组擦痕测量数据划分区域，逐一计算各自不同区域的一阶不变矩，将两组擦痕最相近的特征提取出来，用来计算擦痕之间的相似度，继而给出弹头之间的相似度。

使用一阶矩对擦痕特征的提取以及弹头相似度的判断都是通过计算机编程求解，从而实现了弹头的自动化比对。

关键词：弹痕 自校准 一阶不变矩 相似度 自动比对

参赛队号 1061403

队员姓名 刘 余 张秀平 梁其辉

参赛密码 _____

(由组委会填写)

一、 问题重述

在 77 制式手枪的枪管中有 4 条凸膛线共 8 个棱，分为 4 个主棱和 4 个次棱。弹头从枪管打出时会留下 8 条擦痕，其中 4 条是主棱造成的，4 条是次棱造成的。由于枪弹通过枪管时只旋转了一个很小的角度，擦痕分布成斜线状。

为了判断两个弹头是否出自同一枪支，传统的做法是利用显微镜来观察两个弹头上的擦痕是否很接近。这样的方法显然是效率很低和准确度不够的，并且当犯罪现场的弹头较多时，这样的比较方法显然很困难。而现代高精度的数据采集设备可以给自动对比弹头弹痕提供一些帮助。利用数据采集设备来采集弹头上弹痕的数据信息，然后利用电脑来进行比较，这样可以很好地解决传统做法存在的一些不足。通过电脑比对，判别一个弹头与样本弹头的相似程度，从而确认该两个弹头是否出自于同一枪支。但是，数据的采集常常会遇到精确度不够的问题，这会给判断的结果带来一定的影响。因此，本文需要完成以下几个问题：

1. 如何处理由位置和姿态造成的测量误差，使得两个弹头可以尽量在相同位置和相同姿态的情况下进行比对；
2. 采用怎样的方法来消除由于弹头表面的损伤、锈迹以及痕迹等造成的数据误差和噪声；
3. 对采集到的弹头的次棱数据，提取怎样特征和如何提取来进行弹头的比对；提取数据的哪一部分进行比对效果最好；
4. 利用之前提出的特征提取方法，比较弹痕两两之间的相似度，并给出与每个弹头相似度较高的前五个弹头的文件号。

二、 问题分析

1. 弹头上的擦痕

由于枪弹通过枪管时只旋转了一个很小的角度，弹头上的擦痕分布成斜线状(与圆柱母线有一个夹角)，弹头上留下的每条擦痕就可以近似看成是一条和母线有一个夹角的直线。因此假设中可以把弹头上的擦痕假设为直线来简化问题。弹头上擦痕和母线的俯视图如图 1 所示。

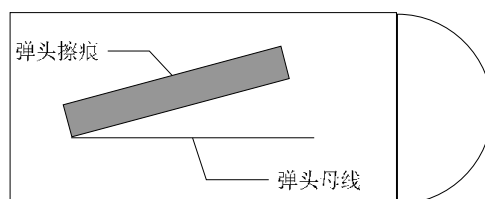


图 1 弹头上擦痕和母线的俯视图

2. 平移误差的来源

弹头最先被固定好在载物台上后，需要移动数据采集设备底部的平移台，使得被测物位于数据采集部件 CCD 下方。每次测量弹头上的弹痕，都必须经过这样的平移操作，

于是给数据的采集带来了平移误差。下面图 2 中的三个图分别说明了在实际值和理想值之间的偏差情况。图中的 xOy 坐标系为基准平面上的坐标系，阴影部分为载物台弹头上的弹痕在基准平面上的投影， Δx 表示是平移误差。

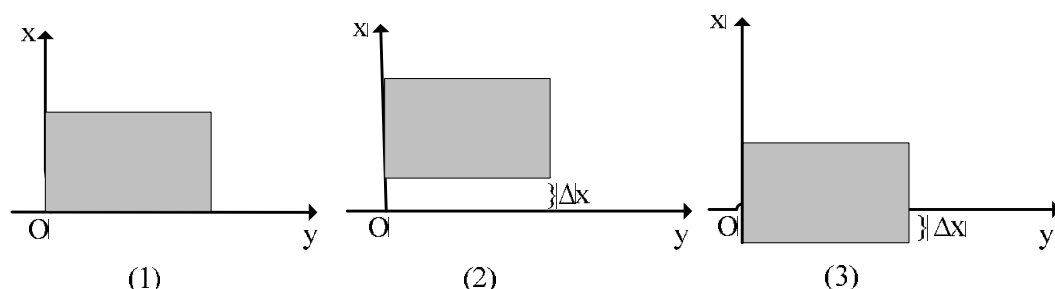


图 2 平移误差来源说明图

下面对图 2 中的 (1)、(2) 和 (3) 分别进行说明。

图 (1)：表示平移弹头使得弹痕能到达最准确的位置，不存在平移误差。但是这只是一种理想的情况。

图 (2)：表示平移弹头使得弹痕超过准确的位置，也就是所测得的 x 的值比实际值要大。

图 (3)：表示平移弹头未使得弹痕达到准确的位置，也就是所测得的 x 的值比实际值要小。

从以上分析可知，完成平移操作过程带来的平移误差，只会对所测得的 x 的值带来影响，而不会对所测得的 y 值和 z 值带来影响。

3. 转动误差的来源

由于在测量弹痕时需要调节弹头的姿态使得：(1) 弹头的圆柱中心线尽量平行于基准平面；(2) y 轴尽量平行于擦痕的走向。而实际上这两个平行都不是准确的平行，他们会带来两方面的转动误差。下面分别对两方面的转动误差对数据测量带来的影响进行说明。

1) 调节弹头使圆柱中心线尽量平行于基准平面带来的误差

从平移台正对着固定好的弹头看（也就是弹头的尾部在左边，头部在右边），为了使弹头的圆柱中心线平行于基准平面，操作的过程是使弹头的某一部（尾部或头部）进行上下摆动。这样的摆动给 y 和 z 的值的测量带来误差，下面通过图例来进行说明。如图 3 所示。

把弹痕投影到 yoZ 平面上，如图中的阴影部分 A 和 B。A 和 B 表示两次不同弹痕的测量时由于转动误差使得他们和 xOy 基准平面的夹角不同。而 A 和 B 上的两点 a 和 b，分别表示弹痕上的一点。

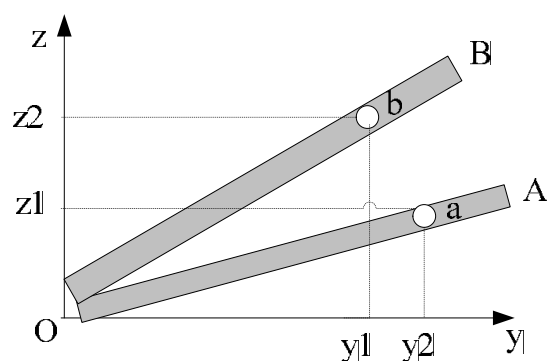


图3 调节圆柱中心线平行于基准平面带来误差的分析

从图中对 a 和 b 两点在 y 轴和 z 轴上的分解可知，调节弹头使圆柱中心线尽量平行于基准平面 xOy 会给测量数据中的 y 值和 z 值带来误差，而没有影响到 x 的测量值。

2) 调节弹头使擦痕的走向尽量平行于 y 轴带来的误差

和上面分析类似，从平移台正对着固定好的弹头看（也就是弹头的尾部在左边，头部在右边），为了使擦痕走向尽量平行于 y 轴，操作的过程是使弹头的某一部（尾部或头部）进行前后摆动。而这样的摆动给 x 和 y 值的测量带来误差，而对 z 值的测量没有带来影响。分析过程示意图如图 4 所示。

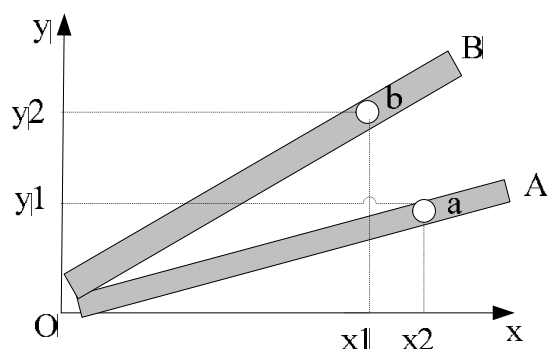


图4 调节擦痕走向尽量平行于 y 轴带来误差的分析

4. 测量数据 x、y 和 z 值的误差来源

根据上面平移误差和转动误差的分析，测量数据 x、y 和 z 值的误差来源总结如下表 1。

表 1 测量数据 x、y 和 z 误差来源

测量数据	误差来源
x	平移误差 + 调节擦痕走向尽量平行于 y 轴带来的误差
y	调节圆柱中心线平行于基准平面带来的误差 + 调节擦痕走向尽量平行于 y 轴带来的误差
z	调节圆柱中心线平行于基准平面带来误差

5. 弹痕数据和图像的关系

每条弹痕的测量数据对应着空间坐标系下的一幅图像，通过对数据的分段可以达到分割图像的目的。这样通过对数据分段的方式，我们把数据处理问题转化为图像处理的问题，利用已有的一些图像分析理论来解决问题。

6. z 值的预处理

在所有的弹头的测试数据中，每一个数据点的 z 值都是保留到小数点后面 6 位数。而根据精度要求， z 的测量精度为 1 微米，数据的单位是毫米。这样，在进行数据处理之前，需要对所有数据的 z 值进行预处理。也就是根据四舍五入的原则，把 z 的值校正到小数点后面 4 位（第 4 位为估算位）。

三、 基本假设

1. 本文中涉及到的枪均指 77 制式的手枪；
2. 假设弹头上的弹痕走向和母线有一个固定的夹角；
3. 假设在调节弹头使圆柱中心线尽量平行于基准平面时，是固定弹头的一端，旋转另一端而实现的；
4. 假设在调节弹头使擦痕的走向尽量平行于 y 轴时，是固定弹头的一端，旋转另一端而实现的；
5. 从同一把枪射出的两颗弹头，弹头的擦痕一定存在相同的特征波形，该特征波形可以用来检验其他弹头是否来自于这把枪；
6. 从不同枪射出的两个弹头，一定不同时存在该特征波形。如果该假设不成立，则就不能利用弹痕特征波形来判断两颗弹头是否来自同一枪支。

四、 弹头位置和姿态调整模型

1. 问题的分析

由于测量条件的差异，各条擦痕的测量数据认为是在不同坐标系下测量得到。而且不论是同一弹头的任意两条擦痕之间，还是不同弹头的任意两条擦痕之间，由于没有一个比对的参照点，所以将两条擦痕的测量数据放到同一坐标系下比对存在一定的难度。解决调整弹头位置和测量误差的一个思路就是利用该弹头自己的测量数据，对自己的位置及姿态进行校准，又称为自校准。

根据问题分析可知，弹头在空间中的姿态是由两个变量决定的。理想情况下，弹头圆柱中心线与基准平面（ xoy 平面）平行，而擦痕的走向与 y 轴平行。但是实际中，很难达到该要求，于是就产生了相应的转动误差。转动误差包括两部分：弹头圆柱中心线与基准平面的夹角 θ 以及擦痕走向与 y 轴的夹角 φ 。调整弹头的姿态就是让 θ 和 φ 尽可能小。

在每次测量时，由于擦痕在数据采集设备 CCD 下所处的位置不同，或者可以简单地认为擦痕没有处于 CCD 的正下方而导致了平移误差。平移误差主要表现数据在 x 轴上坐标的偏移。对于平移误差的修正，同样没有相应的参照物，但是只要求相比较的两个弹头在相同位置下就可以了，所以水平误差可以以一个弹头的 x 轴坐标为参考，沿 x 轴平移另一个弹头的坐标即可。

由于擦痕的尺寸（擦痕宽约为 1.54825mm，长约为 2.07625mm）和弹头的尺寸比较起来比较小，因此忽略擦痕因弹头的圆柱曲面而造成的影响。如果不考虑擦痕底部的高低不同，那么擦痕在外观上可以近似认为是一个长方体。理想状态下该长方体在测量设

备上的空间直角坐标系位置如图 5 所示：

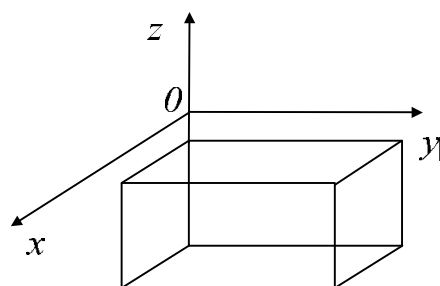


图 5 擦痕在测量设备上的空间直角坐标系理想位置

弹头的空间姿态可以通过擦痕的空间姿态来反映。弹头转动偏差如图 6 所示。为了简便起见，将两个转动偏差 θ 和 φ 分别显示在图 6 的两张子图上。在实际中，子弹在空间中的位置同时包含了这两种偏差。

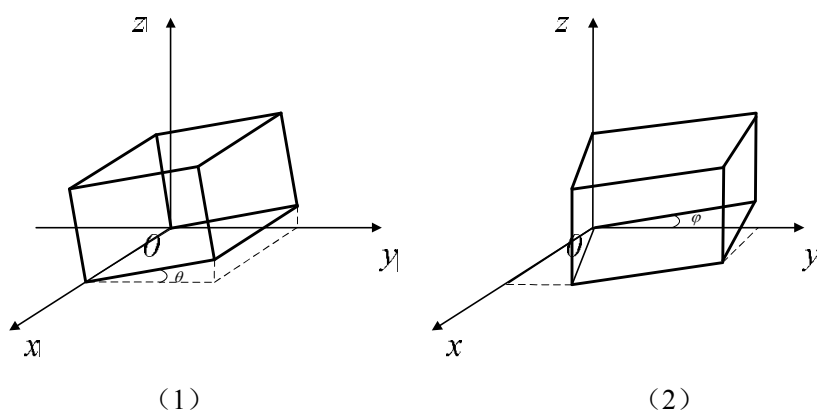


图 6 通过擦痕来表示弹头的空间姿态

(1) 圆柱中心线与基准平面不平行导致的偏差

(2) 擦痕走向与 y 轴不平行导致的偏差

2. 弹头姿态的校正

2.1 圆柱中心线与基准平面不平行导致的偏差 θ 的校正

弹头的空间姿态可以分解为如图 6 所示的两种姿态。这里假设在调整弹头圆柱中心线与基准平面相平行时，是固定弹头的一端，绕 x 轴旋转另一端而实现的。擦痕外轮廓在 xoz 平面上的投影如图 7 所示。

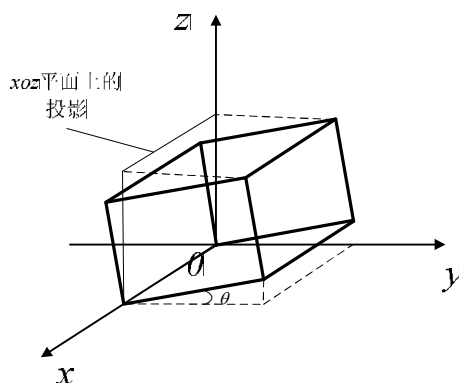


图 7 擦痕外轮廓在 xoz 平面上的投影

从图 7 上可以看出，由于转动误差的存在，造成在 xoz 平面上的投影面积与图 5 在 xoy 平面投影相比会变大。擦痕在 xoz 平面上的实际投影如图 8 所示。图中垂直于 x 轴的直线是当固定 x 坐标时，在 y 方向上采样得到的擦痕高度曲线在 xoz 平面的投影。为了更加直观，只显示了部分数据。转动误差 θ 直接影响了擦痕在 xoz 平面上投影的面积。

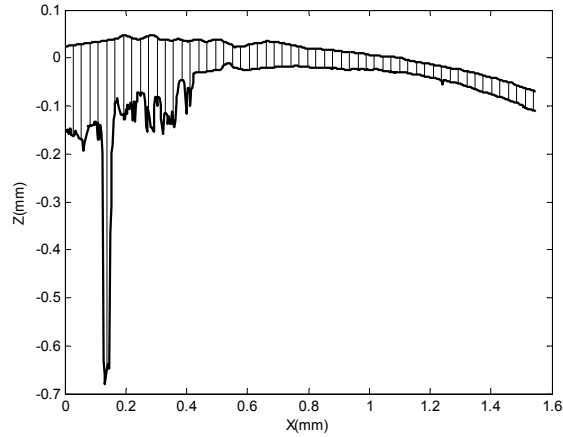


图 8 擦痕在 xoz 平面上的投影

设当固定 x 坐标时，在 y 方向上采样得到的擦痕高度曲线在 xoz 平面的投影的长度为 L ， x 轴上固定第 i 个坐标得到的投影长度为 L_i ， $i=1,2,\dots,N$ (N 为 x 轴上数据的采样点数)，则投影的面积可以表示为：

$$S_1 = \Delta x \cdot \sum_{i=1}^N L_i \quad (1)$$

其中， Δx 表示 x 轴上相邻数据点的距离。

2.2 擦痕走向与 y 轴不平行导致的偏差 φ 的校正

在调节擦痕走向时，同样假设固定弹头的一端，然后通过对另一端绕 z 轴旋转实现的。擦痕走向与 y 轴不平行对数据产生的影响是产生图形的不均匀的压缩。图 9 显示了对一条抛物线做旋转后， yoz 平面上投影产生了不均匀压缩效果。造成此结果的原因是，图形绕 z 轴转动，图像上各个点的 y 坐标的变化量不相同造成的，如图 10 所示。题目中给出的擦痕数据在 xoy 平面也都有不同程度的压缩。为了调整以使得擦痕与 xoy 尽可能平行，应尽可能减小图像的不均匀压缩。即使图 10 中的阴影面积尽可能小。

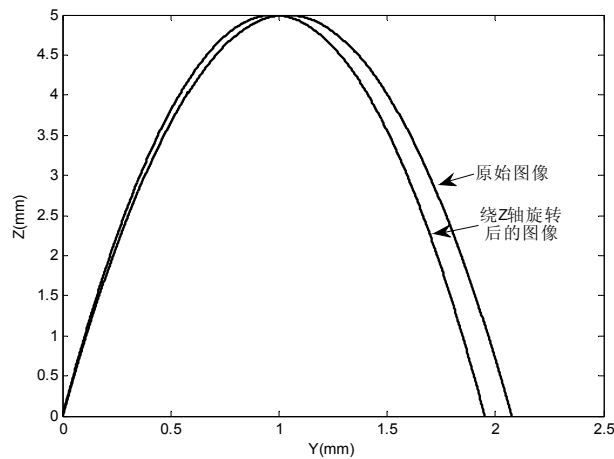


图 9 对抛物线旋转 20° 后产生的图像不均匀压缩效果示意图

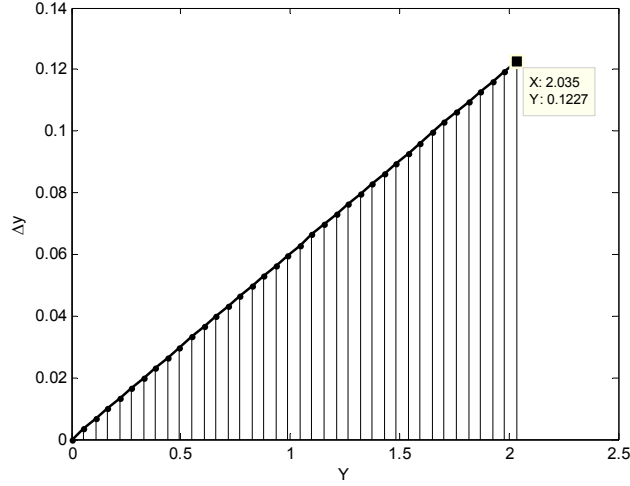


图 10 由于旋转造成的 y 坐标不均匀变化

擦痕测量数据坐标为 y_i , $i=1,2,\dots,N$, N 是测量数据点数。沿 z 轴旋转角度 φ 后, 得到的坐标为 y'_i , 则阴影部分的面积表示为:

$$S_2 = \Delta y \cdot \sum_i^N H_i \quad (2)$$

其中, H_i 为旋转后 y 坐标的变化量, Δy 表示 y 轴上相邻数据点的距离。

2.3 弹头姿态调整算法

结合前面分析的圆柱中心线与基准平面不平行导致的偏差 θ 的校正方法和擦痕走向与 y 轴不平行导致的偏差 φ 的校正方法, 现提出弹头姿态的调整算法。由于没有参照对象, 所以弹头姿态的调整方法属于自校准, 即将该弹头上的待比较擦痕尽可能地调整到与 xoy 平面 (即基准平面) 平行, 而且擦痕的走向与 y 轴平行, 这样就能够保证两颗弹头 (主要是待比较的擦痕) 在相同的姿态下进行比对。

弹头姿态调整算法表述如下:

Step1: 确定调整目标, 即保证待比较的擦痕与基准平面平行, 且擦痕走向与 y 轴平行, 即 $\min S_1 + S_2$;

Step2: 采用搜索的方法搜索转动误差 θ 和 φ , 其操作过程如下:

- 1) 首先给出转动误差 θ 和 φ 的初始变动范围。根据题目中所给出的信息, 转动误差约为 0.2° 。所以初始的转动误差变动范围定为 (但不局限于) $\pm 0.2^\circ$ 。后面可以看到, 随着范围的进一步调整, 误差的变动范围会稍有扩大;
- 2) 设置搜索步长, 在设定误差范围内搜索满足 Step1 中目标条件的转动误差 θ 和 φ ;
- 3) 使用 2) 中得到的结果, 重新设置转动误差的变动范围, 进一步搜索转动误差 θ 和 φ 。以 2) 中得到初步结果为中心, 以几倍于上一步搜索步长的长度为区间长度, 得到新的搜索区间, 然后根据 Step1 中的目标条件, 搜索更加合适的转动误差, 直到满足预设精度为止。

Step3: 利用搜索得到的转动误差, 重新计算擦痕的空间坐标, 从而实现了弹头姿态的自校准。

以第三题数据中编号为 1203959 的手枪发射的两发子弹数据为例，应用弹头姿态调整算法，得到的结果如表 2 和表 3 中的 θ 和 φ 值所示。

2.4 平移误差的修正

两个弹头的平移误差主要是在其放置时， x 轴坐标的差异而导致的。由于没有参照点，所以从擦痕自身测量数据中寻找标记位。通过分析各条擦痕数据，可以知道，数据在 xoz 平面上投影接近于一条抛物线。通过对测量数据进行平均、拟合可以得到一条抛物线。抛物线的对称轴即标记了该抛物线的位置。对于每条擦痕，都有这样一条对称轴来标记其在 x 方向的位置，那么当两颗子弹上面的擦痕进行比对时，可以通过比较其对称轴（数据经过平均、拟合后得到的）位置的差别即可调整平移误差。这种比对是可行的。因为 xoy 基准面是固定在测量设备上的，所以两颗子弹的 x 坐标是对应一个坐标系而言的，从而具有可比性。

以第三题数据中编号为 1203959 的手枪发射的第一发子弹 (T_1) 的第二条擦痕 (C_2) 为例，图 11 给出了该擦痕在 xoz 平面上的投影。通过对测量数据沿 x 轴取平均，即相同 x 坐标的数据取平均值，然后对平均曲线进行二项式拟合，得到的图形如图 12 所示。表 2 和表 3 中给出了第三题数据中编号为 1203959 的手枪所发射的两颗弹头上 8 条擦痕的对称轴坐标 x 。

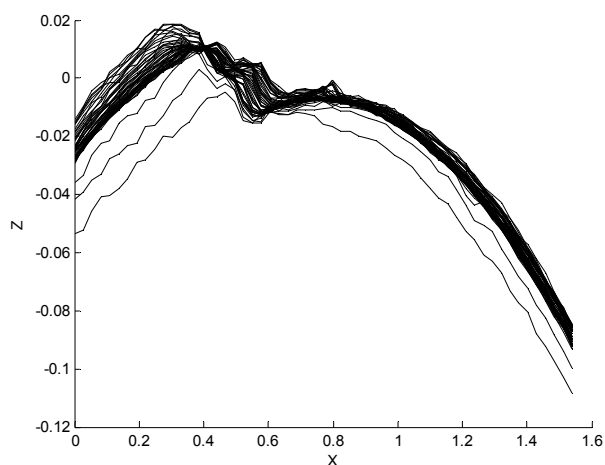


图 11 擦痕在 xoz 平面的投影

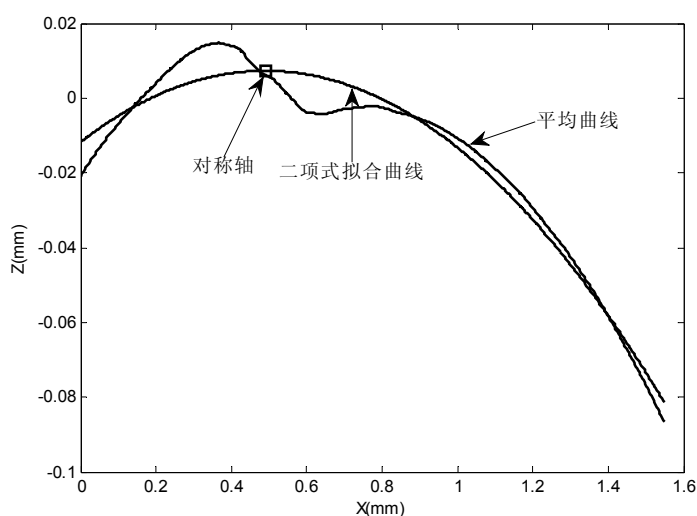


图 12 通过平均、拟合的方法对擦痕进行定位

2.5 弹头位置和姿态的调整结果

以第三题数据中编号为 1203959 的手枪发射的两发子弹数据为例，初始搜索区间设置为 $\pm 0.2^\circ$ ，步长为 15，新区间范围设置为 4 倍上一步搜索步长。两发子弹的位置和姿态的调整方案如表 2 和表 3 所示。

表 2 77T1-1203959 数据的自校正结果

	$\theta (^\circ)$	$\varphi (^\circ)$	$x (mm)$
C_1	-0.2457	-0.0082	0.47575
C_2	0.2285	0.0553	0.49225
C_3	0.2735	0.0456	0.66000
C_4	0.1657	0.0082	0.63800

表 3 77T2-1203959 数据的自校正结果

	$\theta (^\circ)$	$\varphi (^\circ)$	$x (mm)$
C_1	0.1657	-0.0194	0.60500
C_2	0.1818	0.0180	0.30250
C_3	0.1650	-0.0018	0.43725
C_4	-0.2156	-0.0277	0.51150

五、 数据误差和噪声消除模型

1. 问题的分析

在第一问中对测量数据的系统误差进行一定校正，使两颗弹头能够在相同的位置和姿态上进行比对后，还应当考虑数据的随机误差问题。由于弹头表面的损伤、锈迹、油斑、杂质物及痕迹生成的随机性，会给测量数据带来数据误差和噪声。而对随机因素给数据测量带来的误差和噪声，我们分别进行了不同的处理。

1) 消除数据误差，我们考虑的是消除弹痕测量数据图像中的尖峰部分。即对 z 值偏差较大的数据点，利用该点的区域平均值来进行替换，而对 z 值偏差较小的数据点，保持原值。

2) 消除数据的噪声，我们考虑的是对于每个数据点，都利用该点的区域平均值来进行替换，以达到平滑图像的目的。

值得注意的是，上述 1) 和 2) 中所述的利用邻域平均值来进行数据点的替换，是只替换数据点的 z 值而保持数据点原来的 x 和 y 值不变。即分析相邻区域是把图像投影到 xoy 平面上来的。

2. 模型的建立

2.1 空间域平均

图像空间域的平均[1][3]就是对含误差或噪声的原始图像 $f(x, y)$ 的每个像素点取一个邻域 S ，计算 S 中的所有像素灰度级的平均值，作为空间域平均处理后图像 $g(x, y)$ 的像素值。即

$$g(x, y) = \frac{1}{M} \sum_{(x, y) \in S} f(x, y) \quad (3)$$

式中 M 为邻域 S 中的像素点数， S 邻域可取为四邻域或者八邻域。图 13 为 S 邻域为八邻域的情况。

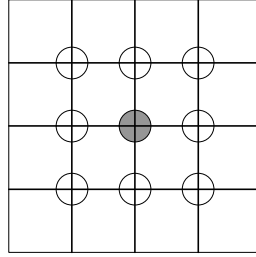


图 13 八邻域局部图

在图 13 中，灰色的数据点表示待校正的数据点，而没有灰色标注的 8 个点为相邻点。校正的过程是首先计算 8 个相邻数据点的平均值，即邻域平均值，然后利用所得的平均值来替换待校正的数据点。

2.2 数据误差消除

对数据误差的消除，我们考虑的是消除数据图像的尖峰部分，也就是对所测得的擦痕的数据中的 z 值的误差进行消除。而数据点校正与否，是通过和阈值的比较来决定的，如下式[1]所示：

$$g(x, y) = \begin{cases} \frac{1}{M} \sum_{(x, y) \in S} f(x, y) & \text{若 } \left| f(x, y) - \frac{1}{M} \sum_{(x, y) \in S} f(x, y) \right| > T \\ f(x, y) & \text{其他} \end{cases} \quad (4)$$

式中 T 为选定的一个非负的阈值，其计算公式如下：

$$T = \frac{|Z_{\max} - Z_{\min}|}{N} \quad (5)$$

式中的 Z_{\max} 和 Z_{\min} 分别表示整个擦痕采集数据中 z 的最大值和最小值， N 表示整个擦痕采集数据的总个数（原始数据中是 40 多万个）。

根据(4)和(5)式，如果某个数据点的值和邻域平均值的差大于阈值 T ，则利用邻域平均值替换该数据点，否则保持原来值不变。

2.3 数据噪声消除

减少噪声常用的方法是进行图像的平滑，一般情况下可以在空间域内采用邻域平均来减少噪声。在频率域，因为噪声频谱多在高频段，因此可以采用各种形式的低通滤波的办法来减少噪声[1][2]。对于测量数据噪声消除问题，我们采用了在空间域进行邻域平均的方法来处理。

也就是对于图像中的每一个像素点 $f(x, y)$ ，采用上面已经给出的式子：

$$g(x,y) = \frac{1}{M} \sum_{(x,y) \in S} f(x,y) \quad (6)$$

来进行校正，以达到平滑整个图像的目的。该处理过程不同于上面数据误差消除，它不需要进行阈值的比较来决定是否校正数据点，而是对每一个数据都用邻域平均值来替换以达到校正的目的。

3. 模型的求解及分析

3.1 数据误差和噪声的消除

以第三问中的 T1-1203959 里的 c1.dat 数据文件为例，数据误差和噪声消除前的原始图像如图 14 所示：

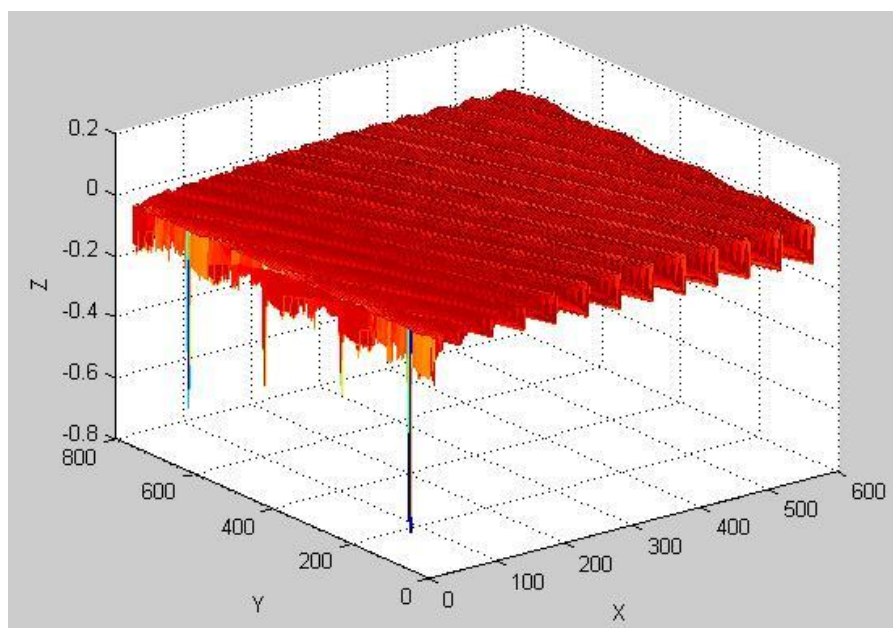


图 14 数据误差和噪声消除前的数据图像

利用上面所述的 2.2 和 2.3 方法来进行数据误差和噪声的消除后，图像如图 15 和图 16 所示：

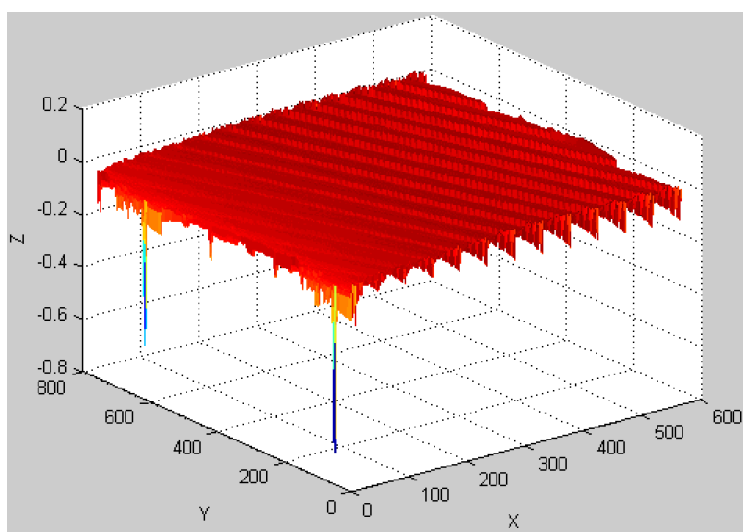


图 15 数据误差消除后的数据图像

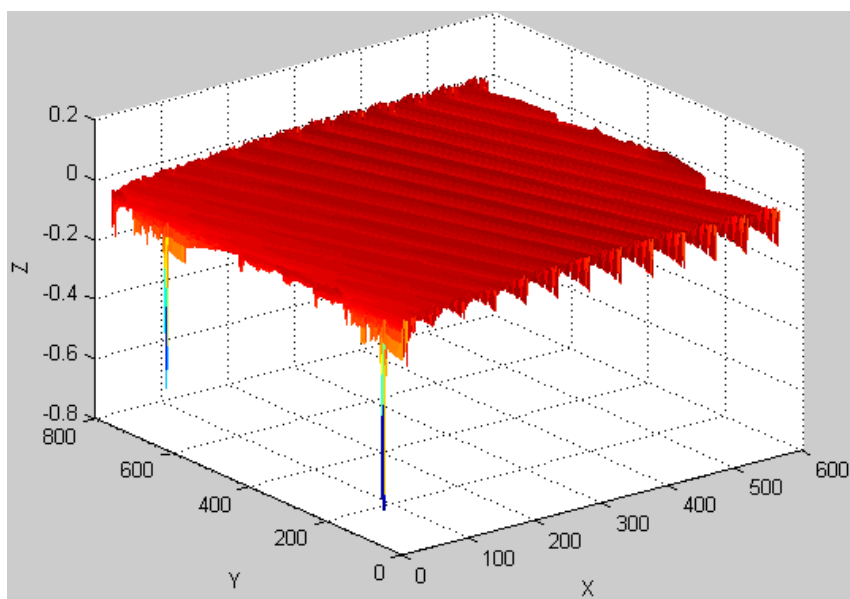


图 16 数据噪声消除后的数据图像

3.2 结果分析

在图 14 的原始数据图像中，图中较多的波峰表明数据的测量存在较大的误差和噪声。经过数据误差的校正后，图像相对于原始图像已经有了很大的平滑；再经过更进一步的噪声消除，数据图像相对于误差消除后的图像要较平滑一些。利用这样的数据误差和噪声的消除方法，在第三问的问题求解中给计算结果的精度带来很大的提高。

值得提出的是，图 14~图 16 三幅图均未能消除边沿数据的误差和噪声，也就是图像中边沿较尖的线条部分。但是由于是边沿数据，并没有给后面的计算带来影响。

4. 模型的检验

以上提出的消除数据误差和噪声上的模型，我们是通过借助第三问中的数据和问题求解的结果来检验的。其主要检验方法是，比较在进行数据误差和噪声消除之后，是否可以达到利用较长的步长来进行数据分段，也可以达到之前采用较短的步长分段数据才能达到的精度。详细的模型检验过程放到第三问的求解中来说明。

六、 数据特征的提取及弹痕比对

1. 数据分析

在第三问所提供的数据中，包含了 6 支枪发射的 12 个弹头（每支枪发射 2 个弹头）的次棱所造成擦痕的测量数据。而每条擦痕的测量数据均以文件的形式保存，每个文件共有 40 多万行，每行有 3 列。每行的 3 个数据值表示擦痕上一个空间点的三维坐标值。而对于同一枪支射出的两个弹头 C 和 D，他们相同编号的擦痕的数据，比如 C_1 和 D_1 ，不一定是由该枪支的同一条次棱造成的，但是 $C_1 \sim C_4$ 和 $D_1 \sim D_4$ 擦痕的数据采集顺序是相同的。也就是说，如果 C_1 和 D_2 同一次棱造成的擦痕，则 C_2 和 D_3 ， C_3 和 D_4 ， C_4 和 D_1 都是由同一次棱造成的擦痕。

2. 几何不变矩

在物体放大、平移、旋转时保持的不变称为不变矩。几何不变矩作为识别的基本特征几经被广泛应用于图像识别、模板匹配及图像分析等邻域中。

对于一幅图像，可以求出该图像的 7 个不变矩。而这些不变矩，当比例因子小于 2 和旋转不超过 45° 的时，对于平移、旋转和比例因子的变化都是不变的，它反映了图像的固有特性。因此，两图像之间的相似度可以利用 7 个不变矩之间的相似性来描述[1]。但是根据对从同一枪支中发射的两个弹头上的擦痕进行分析可得知，只用一个不变矩 I_1 就能够较准确地验证两发子弹是否从同一枪支中发出。

对于离散图像 $f(x, y)$ 来说，矩和中心矩的公式如下：

$$\text{矩:} \quad m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad (7)$$

$$\text{中心矩:} \quad \mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (8)$$

$$\text{式中 } \bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}。$$

根据式 (7) 和 (8)，设置 p 和 q 分别为不同的值，可得 $\mu_{00} = m_{00}$ ， $\mu_{20} = m_{20} - \bar{j}m_{10}$ 和 $\mu_{02} = m_{02} - \bar{k}m_{01}$ 。

为了得到一阶不变矩 I_1 ，再定义一种归一化的中心矩 y_{pq} ：

$$y_{pq} = \frac{\mu_{pq}}{\mu_{00}^r} \quad (9)$$

$$\text{式中 } r = \frac{p+q}{2}, \text{ 从而有}$$

$$I_1 = y_{02} + y_{20} \quad (10)$$

本文中均用这个量来描述子弹擦痕的局部图像特征。

在第三问所提供的数据中，包括了第一、二问中所提到的误差，这给利用特征值进行弹头的准确比对带来了困难。弹头上弹痕采集的数据对应着一幅图像，采用图像的不变矩特征量可以很好地避免这个问题。

3. 问题的分析

根据以上对不变矩的分析，我们可以利用不变矩来对弹头上的弹痕进行比对。对数据进行分段后（即对图像进行区域划分），对每一段数据进行不变矩的求解和比较过程就是特征的提取过程。具体实现过程在下面的问题求解中进行说明。

在不变矩的求解和比较的过程中，同时记录使得不变矩差的绝对值最小的数据段，而该数据段就是所要找的比较效果最好的数据部分。具体过程也在下面问题求解中详细说明。

4. 问题的求解

对每一把枪射出的两颗弹头 C 和 D，对弹头 C 上的每一条擦痕数据 $C_i (1 \leq i \leq 4)$ ，把 C_i 和弹头 D 上的所有擦痕 $D_1 \sim D_4$ 全部进行比对，也就是对出自同一把枪的弹头 C 和 D，共需进行 16 次的擦痕比对。完成这样 16 次擦痕的比对后，就可得知两颗弹头的 8 条擦痕中，哪些擦痕是吻合较好的。

下面以一个简单的例子来说明怎样比较两条擦痕相似度的问题。

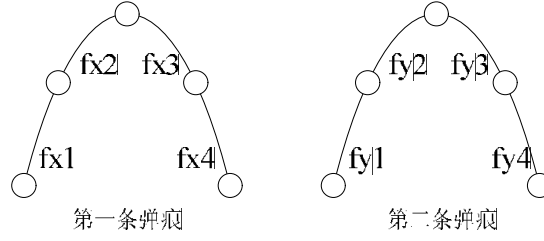


图 17 对两条弹痕进行数据分段

图 17 表示两条弹痕图像在平面上的投影， $x1 \sim x4$ 和 $y1 \sim y4$ 分别表示左右两条弹痕图像被分成的 4 个数据段。利用不变矩公式求得每个数据段的不变矩分别记为 $fx1 \sim fx4$ 和 $fy1 \sim fy4$ 。则两条弹痕的相似度我们定义为：

$$q = \min_{1 \leq i \leq 4} \{ \min_{1 \leq j \leq 4} (|fx_i - fy_j|) \} \quad (11)$$

显然， q 值越小，则两条弹痕的相似度越高。

记第 k 把枪为 G_k ($1 \leq k \leq 4$)，它所发射的两颗弹头为 C 和 D ， $C_1 \sim C_4$ 和 $D_1 \sim D_4$ 分别表示对两颗弹头上各自 4 条擦痕所采集的数据。下面给出问题求解的详细的算法过程：

Step1: 对于任意抢射出的两颗弹头 C 和 D ，均完成以下操作；

Step2: 对于第一颗弹头的每一条擦痕 C_i ，要和第二颗弹头的全部擦痕 D_j ($1 \leq j \leq 4$) 进行比对，其比对的过程如下：

1. 分别对 C_i 和 D_j 按 Δl 的步长进行数据分段，记数据段数都为 N ；
2. 利用不变矩公式分别求出 C_i 和 D_j 的每一个数据段的不变矩；
3. 计算出 $q = \min_{1 \leq i \leq N} \{ \min_{1 \leq j \leq N} (|fc_i - fd_j|) \}$ ，并记录求得 q 值时 C_i 和 D_j 所对应的数据段位置；

Step3: 利用 Step2 中求得的每条擦痕的最小值 q ，再求他们的最小值和该最小值所对应的数据段；

Step4: 利用 Step3 的结果，可以求得两颗弹头 C 和 D 中那两条弹痕是配对的（也就是由同一次梭造成的）。

利用该算法，我们可以确定从同一把枪射出的不同弹头的擦痕间的相似关系。根据经验，弹头的 2 到 3 条擦痕如果吻合较好则可用于鉴定弹头来自同一枪支。也就是说对于同一把枪射出的两个弹头，我们总可以分别找到 2 到 3 条擦痕，它们很相似，从而得到这两个弹头擦痕和擦痕的匹配关系，为后面鉴定两个弹头是否来自同一把枪做准备。

5. 结果及分析

利用上面问题求解中描述的算法，下面以第一把枪发出的两颗弹头上擦痕的数据为例，将搜索到最小 q 值时两颗弹头相应两条痕迹的数据部分画于下图。

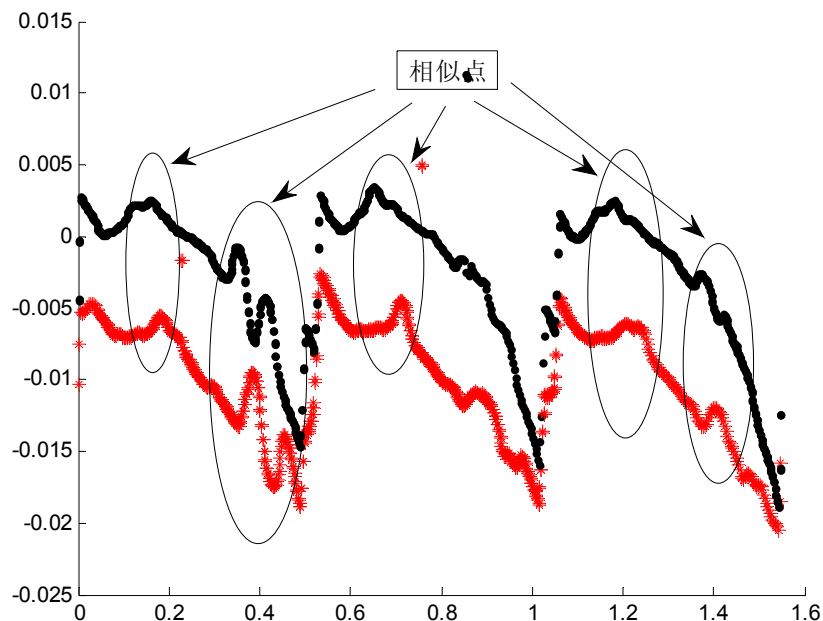


图 18 求得最小值 q 时两颗弹头相应两条痕迹的图像

从图 18 中可以看出，两个图形是十分相似的，而且在由圆圈勾出来的图形细节处，两个图形也吻合的很好。这说明由一阶不变矩来衡量两个图形的特征是合理的。只要一阶不变矩是接近的，那么由此搜索出来的图形就是十分相似的。相对于传统方式而言，现在我们可以用不变矩通过电脑来自动搜索并判别线条的粗细分布是否吻合。

6. 用于消除数据误差和噪声的模型验证

消除数据误差和噪声的前后，数据分段的步长均为 10000，从同一把枪射出的不同子弹擦痕间的相似度见表 4。

表 4 消除噪声前后 q 值计算结果的比较

枪编号	q	
	滤除噪声前	滤除噪声后
1203959	3.39e-3	2.04e-5
1504519	6.06e-4	1.82e-5
1811345	1.12e-4	1.07e-5
1812492	0.53	7.46e-6
1923252	1.34e-05	6.80e-6
1928033	1.27e-3	1.34e-6

从表 4 中可以看出，在滤除噪声后， q 的精度提高了。也就是说，在相同的数据分段步长下，消除噪声更容易提取到图形的特征。这正是因为由锈迹、油斑、杂质物等对原图形造成的非几何系统失真已变小。

七、 22 个弹头擦痕的比较

1. 数据分析

在题目第四问中，22 个弹头上擦痕的测量数据的保存方式和第三问中的一样，共有 88 条弹痕的数据。而对 22 个弹头的痕迹进行两两之间的比较共需 $C_{88}^2 = 3828$ 次比较。

2. 问题分析

在第三问中，我们已经利用不变矩特征来对两条弹痕进行比对，进而求得两条弹痕的吻合程度。在对 22 个弹头痕迹两两之间相似程度的求解时，我们采用相同求解方法。

对于求每个弹头按相似度由高到低给出与之相似的前 5 位弹头问题，可以利用第一小问中求得的弹痕两两之间相似程度的结果，来求得弹头之间的相似程度。

3. 求解过程

3.1 弹痕两两之间的相似度

对于 22 个弹头的弹痕两两之间进行比较的问题，采用的是第三问中的方法。显然这里我们只关心弹痕之间不变矩差的绝对值的最小值，而不再关心哪段数据用于比对效果最好。其主要求解过程为：

- 1) 对于任意的两条弹痕 C 和 D，先按照某一步长 Δl 对测量数据进行分段，分别求出 C 和 D 中每一个数据段的不变矩；
- 2) 利用 C 中的每一个不变矩，和 D 中的所有的不变矩进行做差和取绝对值运算，并记录所有比较后的最小值；
- 3) 在第 2 步求得的所有 C 的数据段保存的最小值中，再取他们的最小值作为两条擦痕的相似度。

设 $fc_i (1 \leq i \leq N)$ 为弹痕 C 第 i 段数据的不变矩， $fd_j (1 \leq j \leq N)$ 为弹痕 D 第 j 段数据的不变矩，其中的 N 为根据某一步长对原始数据（40 多万行）进行划分后的所得的数据总段数。则两条弹痕的相似度可以表示为：

$$q = \min_{1 \leq i \leq N} \{ \min_{1 \leq j \leq N} (|fc_i - fd_j|) \} \quad (12)$$

并且我们定义为弹痕的相似度越高，则 q 值越小。

3.2 弹头的相似度

在求得所有弹痕两两之间的相似程度后，就可以利用弹痕的相似程度来求得弹头两两之间的相似程度。其求解过程如下。

对于任意的两颗弹头 $T1$ 和 $T2$ ，弹头上的 4 条擦痕分别记为 $C_i (1 \leq i \leq N)$ 和 $D_i (1 \leq i \leq N)$ ，则两弹头的 8 条弹痕之间的匹配情况共有四种：

- (I): $C_1 \sim D_1; C_2 \sim D_2; C_3 \sim D_3; C_4 \sim D_4$
- (II): $C_1 \sim D_2; C_2 \sim D_3; C_3 \sim D_4; C_4 \sim D_1$
- (III): $C_1 \sim D_3; C_2 \sim D_4; C_3 \sim D_1; C_4 \sim D_2$
- (IV): $C_1 \sim D_4; C_2 \sim D_1; C_3 \sim D_2; C_4 \sim D_3$

又根据实际情况，如果两个弹头是从同一把枪中射出的，那么这两个弹头上的弹痕

会有 2~3 条弹痕时吻合较好的。按照这样的特点，对于每一种匹配情况，我们算出一个 Q 值，它是 4 对匹配擦痕中两个相似度最大（即 q 值最小）的值的和，也就是：

$$Q = t1 + t2 \quad (13)$$

式中的 $t1$ 和 $t2$ 分别表示每一种匹配情况中的第 1~4 个弹痕对之间的相似度最大的两个值（即 q 值最小的两个值）。

在式（13）的基础上，我们就可以定义弹头之间的相似度，其定义如下：

$$P = \min(Q1, Q2, Q3, Q4) \quad (14)$$

式中的 $Q1, Q2, Q3$ 和 $Q4$ 分别表示利用（13）式分别求四种匹配情况 I、II、III 和 IV 的值。

利用（14）式，就可以求出与每个弹头相似度最高的前 5 个弹头。

4. 求解结果

4.1 22 个弹头痕迹两两之间的相似程度列表

利用 3.1 中提供的方法，可以求得 88 条弹痕两两之间的相似度，共有 $C_{88}^2 = 3828$ 种结果。所求得的相似度的列表很大，故仅从总列表中选取部分放到正文。而全部相似度的列表数据被包含到电子档的附件中。

表 5 22 个弹头痕迹两两之间的相似度部分列表

	1	2	3	4	5	6
1	0.000000	0.000159	0.000003	0.000006	0.000169	0.000352
2	0.000159	0.000000	0.000048	0.000279	0.000002	0.000002
3	0.000003	0.000048	0.000000	0.000031	0.000012	0.000015
4	0.000006	0.000279	0.000031	0.000000	0.000002	0.000001
5	0.000169	0.000002	0.000012	0.000002	0.000000	0.000002
6	0.000352	0.000002	0.000015	0.000001	0.000002	0.000000
7	0.000017	0.000006	0.000014	0.000127	0.000001	0.000007
8	0.000276	0.000645	0.000049	0.000030	0.000386	0.000628
9	0.007295	0.000002	0.001026	0.000102	0.000003	0.000007
10	0.005812	0.000007	0.000000	0.000376	0.000001	0.000007
11	0.000060	0.000000	0.000004	0.000129	0.000014	0.000011
12	0.006005	0.000002	0.000071	0.000010	0.000012	0.000002
13	0.000052	0.000352	0.000187	0.000111	0.000155	0.003095
14	0.000047	0.000007	0.000004	0.000006	0.000004	0.000000
15	0.000016	0.000046	0.000061	0.000124	0.000003	0.000009
16	0.000003	0.000013	0.000004	0.000012	0.000033	0.000114
17	0.010654	0.000007	0.004386	0.000252	0.000001	0.000004
18	0.000040	0.000022	0.000024	0.000001	0.000004	0.000002
19	0.004231	0.000003	0.000011	0.000137	0.000001	0.000002
20	0.000005	0.000028	0.000009	0.000066	0.000036	0.000011
21	0.000016	0.000016	0.000020	0.000013	0.000012	0.000011
22	0.000047	0.000001	0.000080	0.000030	0.000028	0.000001
23	0.010495	0.000000	0.004226	0.000003	0.000116	0.000001

24	0.000304	0.000012	0.000034	0.000330	0.000009	0.000025
25	0.000059	0.000037	0.000012	0.000054	0.000004	0.000007
26	0.000006	0.000004	0.000003	0.000003	0.000021	0.000002
27	0.000000	0.000095	0.000022	0.000021	0.000082	0.000176
28	0.000039	0.000004	0.000036	0.000026	0.000002	0.000004
29	0.000524	0.000074	0.001963	0.000757	0.000335	0.000244
30	0.000000	0.000363	0.000226	0.000005	0.000117	0.010121
31	0.000047	0.000014	0.000027	0.000018	0.000085	0.000025
32	0.000047	0.000014	0.000027	0.000018	0.000085	0.000025
33	0.000278	0.000812	0.000051	0.000222	0.000882	0.000630
34	0.003424	0.000007	0.000015	0.000563	0.000003	0.000001
35	0.000043	0.000001	0.000024	0.000003	0.000005	0.000002
36	0.000175	0.000001	0.000054	0.000151	0.000047	0.000004
37	0.000072	0.000000	0.000010	0.000041	0.000005	0.000001
38	0.000099	0.000309	0.000224	0.000086	0.000015	0.001015
39	0.000073	0.000008	0.000001	0.000026	0.000001	0.000021
40	0.000046	0.000016	0.000021	0.000012	0.000022	0.000000

由于所求得的结果列表较大，正文中只提供了 1~6 条弹痕（在第一行中标出）和 40 条弹痕（在第一列中标出）两两之间的相似度的比较。表中对应两条弹痕的比较值越小，则说明两条弹痕的相似度越高。而由于擦痕间的相似关系是相对的，比如说擦痕对（36，40）和擦痕对（40，36）之间的相似度是相等的；同一条弹痕之间的相似度的值为 0。

上面表 5 中列举的擦痕编号是自定义的简化编号，其对应的实际弹头擦痕文件编号见表 6。表中擦痕文件编号的前 12 个字符代表的是弹头文件编号，c1, c2, c3 和 c4 分别表示每个弹头文件中的四条擦痕数据文件名。

表 6 定义的擦痕编号和实际弹头擦痕的对应关系

编号	擦痕文件编号	编号	擦痕文件编号
1	77T1-1202999-c1	45	77T2-1814117-c1
2	77T1-1202999-c2	46	77T2-1814117-c2
3	77T1-1202999-c3	47	77T2-1814117-c3
4	77T1-1202999-c4	48	77T2-1814117-c4
5	77T2-1202999-c1	49	77T1-1814688-c1
6	77T2-1202999-c2	50	77T1-1814688-c2
7	77T2-1202999-c3	51	77T1-1814688-c3
8	77T2-1202999-c4	52	77T1-1814688-c4
9	77T1-1203959-c1	53	77T2-1814688-c1
10	77T1-1203959-c2	54	77T2-1814688-c2
11	77T1-1203959-c3	55	77T2-1814688-c3
12	77T1-1203959-c4	56	77T2-1814688-c4
13	77T2-1203959-c1	57	77T1-1923252-c1
14	77T2-1203959-c2	58	77T1-1923252-c2
15	77T2-1203959-c3	59	77T1-1923252-c3

16	77T2-1203959-c4	60	77T1-1923252-c4
17	77T1-1504519-c1	61	77T2-1923252-c1
18	77T1-1504519-c2	62	77T2-1923252-c2
19	77T1-1504519-c3	63	77T2-1923252-c3
20	77T1-1504519-c4	64	77T2-1923252-c4
21	77T2-1504519-c1	65	77T1-1928033-c1
22	77T2-1504519-c2	66	77T1-1928033-c2
23	77T2-1504519-c3	67	77T1-1928033-c3
24	77T2-1504519-c4	68	77T1-1928033-c4
25	77t1-1811345-c1	69	77T2-1928033-c1
26	77t1-1811345-c2	70	77T2-1928033-c2
27	77t1-1811345-c3	71	77T2-1928033-c3
28	77t1-1811345-c4	72	77T2-1928033-c4
29	77t2-1811345-c1	73	77T1-1930832-c1
30	77t2-1811345-c2	74	77T1-1930832-c2
31	77t2-1811345-c3	75	77T1-1930832-c3
32	77t2-1811345-c4	76	77T1-1930832-c4
33	77t1-1812492-c1	77	77T2-1930832-c1
34	77t1-1812492-c2	78	77T2-1930832-c2
35	77t1-1812492-c3	79	77T2-1930832-c3
36	77t1-1812492-c4	80	77T2-1930832-c4
37	77t2-1812492-c1	81	77T1-1931817-c1
38	77t2-1812492-c2	82	77T1-1931817-c2
39	77t2-1812492-c3	83	77T1-1931817-c3
40	77t2-1812492-c4	84	77T1-1931817-c4
41	77t1-1814117-c1	85	77T2-1931817-c1
42	77t1-1814117-c2	86	77T2-1931817-c2
43	77t1-1814117-c3	87	77T2-1931817-c3
44	77t1-1814117-c4	88	77T2-1931817-c4

4.2 与每个弹头相似度最高的前 5 个弹头列表

利用 3.2 中提供的方法，可以求得与每个弹头相似度最高的前 5 个弹头列表。求解的部分结果见表 7。

表 7 与每个弹头相似度最高的前 5 个弹头列表

	第 1	第 2	第 3	第 4	第 5
1	3	21	7	16	5
2	3	5	16	10	15
3	2	15	1	18	6
4	20	3	6	5	1
5	18	3	2	16	4
6	3	15	16	7	4
7	18	17	3	16	6

8	9	13	21	7	17
9	12	3	17	7	16
10	17	2	15	20	11
11	17	18	10	22	1
12	9	20	18	22	14
13	15	16	14	1	8
14	22	18	15	13	1
15	3	6	10	2	20
16	17	2	6	3	7
17	7	22	16	18	10
18	7	5	17	3	11
19	16	3	4	2	20
20	12	3	4	15	10
21	1	16	4	18	8
22	17	3	14	20	11

表 7 中的第一列数据表示待判断的每一颗弹头的标号，第一行表示相似度由高到低给出的 5 个弹头的编号。对于每一行数据值，从左到右弹头的相似度是递减的。

同样，表 7 中的弹头编号只是自定义的简化编号，实际的弹头文件编号对应于表 8 中。

表 8 表 7 中自定义的弹头编号对应的弹头文件编号

编号	弹头文件编号	编号	弹头文件编号
1	77T1-1202999	12	77T2-1814117
2	77T2-1202999	13	77T1-1814688
3	77T1-1203959	14	77T2-1814688
4	77T2-1203959	15	77T1-1923252
5	77T1-1504519	16	77T2-1923252
6	77T2-1504519	17	77T1-1928033
7	77t1-1811345	18	77T2-1928033
8	77t2-1811345	19	77T1-1930832
9	77t1-1812492	20	77T2-1930832
10	77t2-1812492	21	77T1-1931817
11	77t1-1814117	22	77T2-1931817

5. 结果分析

在表 7 中，对于子弹 7，子弹 18 和 17 与它最相似。对于子弹 17，和它最相似的是子弹 7。这这说明它们的相似关系具有对称性。像这样相似关系具有对称性的子弹对还有 (5, 11)，(10, 15)，(12, 20)。值得注意的是，并不是所有的子弹对都满足这种对称关系。但是对于子弹对 (i, j) ，若子弹 j 是与子弹 i 最相似的 5 发子弹之一，那么反之也成立。这正说明即使是从同一把枪射出的不同子弹，由不变矩衡量的子弹间相似关系也是有一个浮动范围。如果它们的相似关系在某个范围之内，则可认为它们是从同一把枪射出的子弹；如果它们的相似关系在该范围之外，则可认为它们不是从同一把枪射出的。

6. 怎样判断两颗子弹出自同一把枪

也就是确定从同一把枪射出的任两颗子弹相似关系的取值范围。该问题其实是寻找一个阈值 T_h ，当两个子弹的相似程度大于 T_h 时，则认为它们来自于同一把枪。反之，则认定它们不是来自于同一把枪。

为了确定该阈值，我们先分析从同一把枪射出的不同子弹之间相似程度的取值范围，记为 $(smin, smax)$ 。而对于不同枪射出的子弹，任两颗子弹相似程度也有个取值范围，记为 $(dmin, dmax)$ 。根据本文对任两颗子弹的相似程度的衡量，采用第 4 问的所有数据进行统计，得出 $smax = 1.73 \times 10^{-3}$, $dmin = 6.25 \times 10^{-4}$ 。从此可以看出，两个区间并没有交叠，并且这两个值相差一个数量级。这也就是说，以一阶不变矩为基础来衡量两个子弹相似关系，效果是明显的。

如果我们把阈值定为 $T_h = (smax + dmin) / 2$ ，则我们可轻易判断出两个子弹是否来自于同一把枪。当两个子弹的相似度量值大于 T_h 时，它们不是来自于同一把枪；当两个子弹的相似度量值小于 T_h 时，它们来自于同一把枪。注意，本文的相似度量值并未采取归一化处理，若相似度量值越小，则说明两个子弹相似程度越高，两个子弹越可能来自于同一把枪；若相似度量值越大，则说明两个子弹相似程度偏差越大，相似程度越低，两个子弹越不可能来自于同一把枪。

八、 模型评价及改进

1. 自校准模型

采用自校准方法调整弹头的位置和空间姿态是一种绝对校准的方法。该方法虽然没有同其他参考系进行比较，但同样使弹头的姿态进行了调整，并且可以使两个不同的弹头尽可能地在同一姿态和位置下进行比对。使用自校准方法得到的结果在很大程度上取决于第一次搜索的步长和后续过程中搜索范围的确定。在第一搜索中，首先确定了最佳调整角度可能存在的位置，然后在该位置在其周围空间中进一步搜索。如果第一次搜索步长取的较大，而后面的搜索范围设置较小的话，很容易陷入局部搜索。也就是说，得到的结果并不是调整弹头姿态的最佳结果。

2. 不变矩特征量

一阶不变矩可以用来衡量图形的特征，从而为比较不同波形的特征以及它们的相似性做出判断。它可以应用于需要比对波形的场合，比如说公安系统中弹痕的比较。但是，一阶不变矩对图形特征的描述是有限的，所以可以考虑用更高阶的不变矩或它们的组合来进一步描述图形特征或更细节的东西。随着更高阶矩的使用，图形特征将会刻画得更仔细，波形的比对将会更准确。

九、 参考文献

- [1] 夏正良，数字图像处理，南京：东南大学出版社，1999
- [2] 沈庭芝等，数字图像处理及模式识别，北京：北京理工大学出版社，1998
- [3] 于殿泓，图像检测与处理技术，西安：西安电子科技大学出版社，2006
- [4] 王炳成，线条状痕迹检验识别中的非线性理论与应用研究，东北大学博士学位论文，

48-51, 2004

- [5] 张洪刚等, 图像处理与识别, 北京: 北京邮电大学出版社, 2006
- [6] 周渭, 测试与计量技术基础, 西安: 西安电子科技大学出版社, 2004
- [7] 薛定宇等, 高等应用数学问题的 Matlab 求解, 北京: 清华大学出版, 2004